

# Investigations into the opportunity for spatial management of the quality and quantity of production in UK potato systems

James A. Taylor<sup>1</sup>, Hongyan Chen<sup>1</sup>, Matthew Smallwood<sup>2</sup> and Bruce Marshall<sup>3</sup>

<sup>1</sup>School of Natural and Environmental Sciences, Newcastle University, Cockle Park Farm, Ulgham, NE61 3EB, UK

<sup>2</sup>McCain Foods GB Limited, Scarborough, YO11 3BS, UK

<sup>3</sup>James Hutton Limited, Invergowrie, DD2 5DA, UK

E-mail: [james.taylor6@newcastle.ac.uk](mailto:james.taylor6@newcastle.ac.uk)

**Keywords:** variograms; tuber size distribution; marketable yield; yield monitor.

**Acknowledgements:** The authors are grateful to the commercial potato growers who provided access to their fields for the survey. The assistance of technical staff at Newcastle University and McCain Food GB Ltd., particularly Mr M. Botha and Mr B. Chapman, during the manual digs and grading is also grateful acknowledged. This project is funded through the Innovate UK Agri-Tech Catalyst Fund in conjunction with the Biotechnology and Biological Sciences Research Council and is supported by SoilEssentials Ltd, McCain Foods GB Ltd and Grimme UK Ltd.

## Abstract:

Effectively adopting site-specific crop management is dependent on there being production variation within the system. There has been very little published material on yield variation in potatoes and almost none on tuber quality. A review of literature and a systematic survey of 13 potato fields in the UK was performed to better understand within-field variability in potato quantity and quality attributes, including stem density, total yield, marketable yield and tuber size distribution. Non-spatial and spatial statistical analysis of manual and sensor observations was performed on the survey data. Ware production fields exhibited more structured spatial variation in quality attributes (tuber size distribution and % marketable yield) rather than yield attributes. Seed production systems were inversed, with more structured spatial variability in yield attributes than quality attributes. Manual and sensor yield data exhibited a similar level of spatial variation, however the spatially denser sensor data indicated a nested effect with a large amount of short-range variation that may be management related. Both quality and quantity parameters showed sufficient magnitude and range in variation for site-specific management to be considered. The results presented here provide a baseline for spatial variation in both potato quantity and quality to inform future work, particularly for work considering spatial management or spatial modelling in potatoes.

## Introduction:

The spatial variability in quantity and quality production attributes is a driver for adoption of site-specific management (Precision Agriculture). Variability in production should exhibit a sufficient magnitude to warrant differential management and this variability should have a sufficient spatial structure to permit variable-rate management (Pringle et al. 2003). The magnitude and spatial structure required in production variability to site-specifically manage a crop will depend on the value of the crop and on the precision/size of farm machinery used (Tisseyre and McBratney 2008). Theoretically, higher value crops have a greater opportunity for differential management as there is a greater potential benefit to profitability with small increases in quality and/or quantity gains. Higher value crops also tend to be managed more intensively. For example, vineyards operate with 2-4 row mechanisation (5-10 m swaths) while cereals operate on 24+ m swaths for most operations. In terms of the value of production (per hectare), field horticulture tends to fall between perennial horticulture (higher value) and annual cereal/legume (lower value) production.

Potato (*Solanum tuberosum*) production is one of the main field horticulture crops produced worldwide, and in the UK is considered a high value crop for arable producers within a cereal/legume/brassica rotation. It seems to be an ideal crop for site-specific crop management and studies into precision potato production have been published since the mid-1990s (Schneider et al. 1997 and 1996; Hess et al. 1998; Persson 1998); however adoption of precision agriculture into potato systems has been slow compared to other annual arable systems. There are multiple potential reasons for this including (among others) a typical 1 in 6 (or greater) rotation for pest and disease considerations in the UK<sup>1</sup> that generates a discontinuity in data collection, the lack of a robust yield monitoring system for potato harvesters (Davenport et al., 2002), the slow development of variable-rate potato specific machinery (Kempenaar et al. 2018), a lack of a spatial decision support structure for potatoes, a current low ability to vary irrigation spatially in fields, and historically good profit margins that mask production inefficiencies. A change in this last reason, with decreasing profit margins within the UK potato industry over the past decade, and improving agri-technologies has led to renewed interest in site-specific potato management to improve production efficiencies. It has also prompted more work in understanding site-specific management options (e.g. Allaire et al. 2014, Cambouris et al. 2014, Whelan and Mulcahy 2016 and 2017, Al-Gaadi et al. 2016, Holmes and Jiang 2017) and modelling the potential profitability of precision potato management (van Evert et al., 2017).

As with most high-value crops, the value in potato production lies not just in the quantity of production but also in the quality. Discounting blemishes, rots or other disease effects, quality is usually associated with the size and shape of the potato and dry matter % in relation to the target market. Potato size is critical for its end use – seed potatoes must fit within a certain size band and likewise ware potatoes also have an ideal size range for the final product (chips, crisps, etc...). Premiums are therefore high to ensure potatoes are delivered to size. Ensuring that as much of the crop as possible fits into the desired market size is one way that growers can improve profitability. There is an effective difference between the total yield and the marketable yield.

To determine if, and to what level, site-specific crop management is applicable in potato production, information on the spatial variability in the quantity and quality of production is needed. To help

---

<sup>1</sup> Guidelines provided by Red Tractor Assurance for Farms – Crop-specific Module: Potatoes (<https://assurance.redtractor.org.uk/contentfiles/Farmers-6609.pdf> - June 2018)

address this need i) a brief review of the literature will be presented that includes reported within-field variance in potato production, ii) results pertaining to the spatial variability in yield and quality parameters observed from a systematic manual dig survey in UK potato fields will be reported and, iii) the yield variance recorded from on-harvester yield sensors within some of the same fields will also be reported. All this information will then be used to identify and discuss the potential opportunity for in-field site-specific potato management that the observed variability imparts.

## **A brief review of spatial variability in potato production**

As with other annual cropping systems, within-field yield variation in potato production is known to occur and has been observed in several studies. The method of reporting this variation has been *ad hoc* but reported statistics on yield variability from manual digs include;

- a mean of 9.01 Mg ha<sup>-1</sup> (as dry matter) and reported a coefficient of yield variation of 10.4% from 36 plot digs (4.5 m<sup>2</sup>/plot) in a Dutch field (Finke et al. 1992),
- a range of 20-50 Mg ha<sup>-1</sup> from 65 small plot digs (plot size not specified) in a Dutch field (Verhagen 1997 and Verhagen et al. 1997),
- a range of 23-59 Mg ha<sup>-1</sup> and 22-48 Mg ha<sup>-1</sup> from 34 samples digs (3 m lengths across 4 rows) in each of two fields in Maine, USA (Starr 2005),
- mean yields and coefficient of variation (CV) ranges of 36.8-43.0 Mg ha<sup>-1</sup> and 8.3-14.5% respectively from 66 sites within a Canadian field in 4 different years (1991, 1992, 1994, 1997 with plots of 1.22 m of row in 1991 and 3.20 m of row in other years). (Rees et al. 2007),
- a total yield CV of 30% and 31% for two fields in Canada, each with 108 sample site digs (1.5 m of row) (Allaire et al. 2014), and
- a mean yield and CV range of 34.1-42.5 Mg ha<sup>-1</sup> and 15.3-36.1% respectively from 45 sites (3 m<sup>2</sup> plot) in each of 3 irrigated fields in Saudi Arabia (Al-Gaadi et al. 2016)

In addition to manual dig observations, yield variance recorded using on-harvester potato yield monitors has been sparsely reported. For three whole fields a CV range of 24.9 – 28.2 % was reported from Canadian studies (Cambouris et al. 2006), while a Swedish study reported CVs of 35% and 19% for two fields (Persson et al. 2005). A more recent survey of 22 fields in Australia reported CVs of yield from 21 – 34% (Whelan and Mulchay 2016).

While relatively few in number, especially when compared to studies on cereal yield, these manual and sensor-based total yield statistics clearly indicate a large magnitude of variation in yield within individual potato fields. Reported CVs of 8-31% are similar to those reported in cereal studies (e.g. Pringle et al. 2003) where site-specific crop management has been adopted. As indicated above, yield variability is usually reported as either a range or as a CV statistic. It is important to emphasise at this point that the range and the CV statistic are non-spatial statistics. They do not account for the spatial area over which the data were collected nor does they give any indication of whether the observed variation is random in nature or shows spatial patterning and trends.

Only three studies were found that performed a geostatistical analysis of yield. Cambouris et al. (2006) and Whelan and Mulcahy (2016) presented spatial information associated with yield monitor data

while Allaire et al. (2014) reported on manual dig data. The Allaire et al. (2014) data is difficult to interpret as it is a grid survey (5-10 m) performed on very small fields (<0.9 ha) and presented as experimental variograms with very little detail. The Cambouris et al. (2006) data indicates that the yield variation observed in the field had a moderate spatial structure but relatively short variogram ranges of 21-36 m. Whelan and Mulcahy (2016) graphically depicted yield variograms from 22 fields but only reported parameters for a derived median variogram. Visually most of the field-specific variograms again had a moderate spatial structure and a variogram range between 15 – 40 m.

Two separate crop simulation exercises with potato yield from the Netherlands generated slightly higher variogram ranges of 47-59 m (Verhagen 1997, Van Uffelen et al. 1997). These observations and simulation results indicate that the variogram range in potatoes may be considerably less than that observed in cereal studies (Pringle et al. 2003). If this is found to be the norm, a consequence of this will be that potato machinery must be able to react more quickly to changes in crop conditions. The reason for the shorter variogram range is not evident and may be a factor related to the small sample size available (3 studies) not being a good representation of expected variability.

The studies reported above have all focussed on total yield variance, not marketable yield variance. Very little information has been reported in the scientific literature on the spatial variability in marketable yield or other quality attributes in potato. The early study from Verhagen et al. (1997) did observe that the 30 Mg ha<sup>-1</sup> yield digs only yielded 3 tons of highly desired potatoes (> 50 mm) (or 10% of total yield) while the 45 Mg ha<sup>-1</sup> plots yielded up to 15 Mg ha<sup>-1</sup> of tubers > 50 mm (30% of total yield). Davenport and Hattendorf (2000) presented maps of specific gravity from bin samples and site samples (~60 sites/field) with some visual spatial patterning but no statistics were reported. Anecdotally, Hetzroni et al. (2007) reported a large variation in market quality potatoes between bins being delivered from the same field but were unable to map or spatially quantify this variation. A recent paper, has reported the percent marketable yield (45 – 80 mm tubers) from different potential management zones within two New Zealand fields (Holmes and Jiang 2017). From 1.5 m manual digs within different zones, Holmes and Jiang (2017) reported within-field differences of 58 – 71 % and 69 – 88 % for the two fields respectively. However, no direct spatial analysis was performed of the tuber size distribution, only a validation of soil-sensor derived zones. Spatial variation in potato quality attributes (size, shape, dry matter %) appears to be an area where more quantifiable data is needed to inform scientific analysis and industry decision-making.

Based on the limited non-spatial and spatial statistics reported in the literature to date, there appears to be sufficient variability in total yield in potato production to encourage site-specific crop management, but no data are available on how this translates through to the marketable yield or to potential site-specific management to maximise marketable yield. The objective of the second part of this paper is to present empirical evidence on spatial production variability from intensive field surveys in both ware and seed production systems, with the intent to start to fill this knowledge gap and to inform future spatial modelling and decision support development in potato systems.

## **Materials and Methods**

A systematic survey of 13 fields was conducted in 2015 and 2016 to gain a better understanding of the spatial variation in potato quality and quantity in UK seed potato (5 fields) and ware potato (8 fields) production systems. Surveyed fields were located in various UK potato regions including

Staffordshire, Yorkshire, Leicestershire and Angus to provide a range of UK production systems. Ware producers were all long-term contractors to McCains Potatoes GB and were selected as having demonstrated best practice operations and considered good collaborators by McCains Potatoes GB. Seed potato for the ware fields was sourced from McCains Potatoes GB. Seed from the same batch was grown in the same location and stored in the same cold store prior to planting. This seed was certified by the Scottish seed inspectorate and additional quality assessments were made by McCain to verify its health status. The seed potato fields belong to a single farmer who was actively involved in the project. Seed potato was stored on-farm from the previous season's production. Potato agronomy in all fields was performed using current best practices. There were no major diseases or pest issues in the fields surveyed.

All fields were either surveyed pre-planting using an Electro-Magnetic Induction (EMI) sensor to generate apparent soil electrical conductivity ( $EC_a$ ) maps of the field or had existing  $EC_a$  maps available. The  $EC_a$  maps were used as a base layer to stratify 100 sample sites within each field to ensure confidence in the generation of the experimental variogram (Webster and Oliver 1992). As part of the stratification process, a minimum of 5 sites in each field were selected to be paired sites, i.e. another unique sampling point was located within 5 m of the original sample location. In this way data existed to understand the amount of variation between sample points that are located close together.

Each sample site was geo-referenced with a WAAS-enabled Garmin eTrek Global Navigation Satellite System receiver (Garmin Ltd, Olathe, KS, USA) and Eastings and Northings (UTM WGS84 Zone 30N) recorded. At each site, a 3 m harvest dig was performed along one bed (all fields were on a double bed set-up) at least 2 weeks after haulm destruction and usually immediately before mechanical harvesting. Both plant and stem numbers were counted before the plants were dug up. Stem density ( $S_D$ ; counts  $ha^{-1}$ ) was preferred to plant density for the analysis as it provided a more reliable measure of the amount of vegetative development at a site. Theoretically, with a well performing planter, plant density should be uniform, but plant numbers were not uniform in the 3 m plots, which is likely due to variability in planter operation. Care was taken to retrieve all tubers from the sample site, including <15 mm tubers, when doing manual digs. Tubers were stored in paper bags and put through grading riddles to generate a tuber size distribution in either 5 mm or 10 mm fractions. Each fraction was also counted and weighed to give tuber number and mass for each 5 or 10 mm tuber size fraction. Using results from tubers > 15 mm, the number of tubers ( $TN_{15}$ , 1000s tubers  $ha^{-1}$ ), the total yield ( $Y_T$ , Mg  $ha^{-1}$ ) and the marketable yield ( $Y_M$ , Mg  $ha^{-1}$ ) was calculated. By dividing  $Y_M$  by  $Y_T$  the % marketable yield (% $Y_M$ ) was determined. The ratio of mass of tubers (total yield) to the number of tubers ( $MNR_{15}$ ) for grades (tubers) greater than 15 mm indicates mean tuber size at a site and has been shown to be a good indicator of tuber size distribution in potato systems (Sands and Regal 1983, Mackerron et al. 1988). The  $MNR_{15}$  was calculated from the graded tuber data and expressed in units of g tuber $^{-1}$ . For each field the marketable yield depended on the variety and target market. Information on riddle size used and the desirable tuber range for each field is given in Table 1.

The collection of yield monitor data, using the RiteYield potato yield monitor (Greentronics, Elmina ON, Canada), was attempted in all ware production fields in both 2015 and 2016. Unfortunately the 2015 harvest was a wet harvest and generated a lot of unusable yield sensor data due to difficulties in separating tubers and soil during harvest. The load cell-based yield monitor was therefore affected by the mass of soil being simultaneously harvested. Yield sensor data is restricted to 2016 ware fields.

These data were trimmed using the protocols outlined in Taylor et al. (2007) to remove extreme values and outliers prior to analysis.

#### *Non-spatial statistical analysis*

Classical statistics ( $\mu$ ,  $\sigma$ , CV, range, skewness ( $\gamma_1$ ) and N) were generated for each field and for the entire data set split along production lines (ware vs. seed) for total yield ( $Y_T$ ), marketable yield ( $Y_M$ ), % marketable yield ( $\%Y_M$ ), mean tuber mass ( $MNR_{15}$ ) and harvest stem density ( $S_D$ ). Ranges across the fields are presented here, rather than individual field values. The split along production lines was done in recognition that the two systems are inherently different in the target market and exhibited differences in management.

<<INSERT TABLE 1 NEAR HERE>>

#### *Geostatistical analysis*

Geostatistical analysis was based on variography. The skewness analysis did not indicate any strongly skewed data, especially within individual fields (data not shown), so no transformation was performed prior to the generation of the experimental variogram. Similarly, data were not normalised, even though variance differed greatly between variables, as the intent here is to present the variance associated with the actual data.

Variography is a method of representing the amount of stochastic variation in the data, known as the nugget ( $c_0$ ), the amount of variance that exhibits some level of autocorrelation, referred to as the sill ( $c_1$ ) and the distance over which this autocorrelation occurs, the range ( $a$ ). This is achieved by calculating the (semi-)variance between pairs of points separated by fixed distances e.g. all pairs of points 5 m apart, 10 m apart, 15 m apart etc. These distances are referred to as lags. The semivariance at each lag is plotted to generate an experimental variogram cloud (Fig. 1). A model is then fitted to the data to describe the shape of the experimental variogram with parameters describing the nugget, sill and range. For further information in a precision agriculture context on the calculation of the variogram and common models that are used to fit the experimental variogram readers are directed to Pringle et al. (2003).

<<INSERT FIGURE 1 NEAR HERE>>

Variogram analysis was performed separately on the global ware and global seed production data sets. The shareware Vesper (Minasny et al. 2005) was used to generate an 'average' variogram for each variable in each type of production system. McBratney and Pringle (1999) used a transformation process to generate 'average' variograms from individual survey (or field) data after a meta-analysis of variograms of soil properties from a wide variety of sources. This was done due to concerns with potentially skewed (non-normal) data. However, since the data sets in this survey have similar and known sampling densities, and were normally distributed (see Results), this approach has not been used, and the data simply concatenated into a single file before variogram analysis. The assumption is

that the data from each field is providing a similar weighting to the variance calculation at each lag within the combined analysis.

Experimental variograms were generated from the concatenated data using 40 lags constrained to a maximum distance of 400m. A spherical model was the best fit using the Akaike Information Criteria to the majority of the experimental variograms (data not shown), so this model was used to fit all 'average' variograms. In the case of the seed production fields, some fields were located adjacent to each other, such that pairs of points at some lags could be calculated with data from different fields. An adjusted Northings was generated to ensure a displacement of at least 500 m between adjacent fields. This eliminated any potential between-field variance contributing to the 'average' variogram analysis but maintains the within-field variance structures. No ware fields were located within 400 m of each other. Variogram parameters ( $c_0$ ,  $c_1$ ,  $a$ ) were recorded for each 'average' variogram. Plots of standardised variograms were also generated for each variable along production lines by dividing the lag semivariance by the total variance ( $c_0 + c_1$ ), which standardised the semivariance to a value between 0 – 1 for all variables.

In the case of the yield sensor data, a global 'average' yield variogram from the sensor data was generated in Vesper using the same approach as for the manual dig data (Spherical model, constrained to 400 m with 40 lags) and the variogram parameters recorded. To avoid bias from different size data sets, the data from each field was randomly subset to 10,000 points and then combined to generate the data set from which the experimental variogram was derived. As there were only limited fields with yield data (5) and there is a lack of reported field-level potato yield monitor data in the literature, the field-specific variograms were also generated and parameters recorded. There were 8 fields without yield data due to a combination of factors, including logistical issues of getting a harvester with a yield monitor into the field and wet soil conditions during harvest causing excessive amounts of soil to pass over the yield sensor and confounding the signal.

The 'average' variogram parameters were used to calculate estimates of spatial structure, namely the Cambardella Index (CI - Cambardella et al. 1994) (Eqn. 1) and the mean correlation index (MCD - Han et al 1994) (Eqn. 2). The normal CV value for each variable from the aggregated data was also generated as a comparison. Lower CI values and larger MCD values tend to be indicative of more spatially structured variation in the data.

Cambardella Index (CI)

$$CI = \frac{c_0}{c_1 + c_0} \bullet 100 \quad \text{Eqn. [1]}$$

where  $c_0$  = nugget and  $c_1$  = sill,

and <25 = Strong spatial dependency

25–75 = Moderate spatial dependency

>75 = Weak spatial dependency

The Mean Correlation Distance (MCD)

$$MCD (m) = \frac{3}{8} \bullet \frac{c_1}{c_0 + c_1} a \quad \text{Eqn. [2]}$$

where  $c_0$  = nugget,  $c_1$  = sill and  $a$  = range

## Results

### *Non spatial statistics*

The range of observed statistical values for the various production attributes across the 8 ware fields and the 5 seed fields are presented in Tables 2 and 3 respectively. The range of field level total yield ( $Y_T$ ) values for both seed and ware were similar, though slightly larger in ware systems, and corresponded with similarly reported studies (Verhagen 1997, Starr 2005, Rees et al. 2007). The CV range for  $Y_T$  was in the same range as those reported previously, though generally in the lower end of this range with the maximum here 17.9% compared to some studies reporting values above 30% and many with values in the 10-20% range.

Marketable Yield ( $Y_M$ ) was more variable than  $Y_T$  (higher CV range) for ware but very similar in variability to  $Y_T$  in the seed systems. The range of percent marketable yield ( $\%Y_M$ ) values was larger in ware systems than in seed systems, which may reflect the greater diversity in growers and regions for the ware systems (vs. only one grower and region for seed). In ware systems, a considerable amount of the crop was not of sufficient size for marketing ( $\%Y_M$  63.17 – 88.61%), with even the best performed system having 11% of produce outside the marketable size.

Stem density ( $S_D$ ) also exhibited considerable variation (CV 12.40 – 31.49%). The mean  $S_D$  was higher in seed systems, as expected with closer tuber planting densities, but the amount of variation was higher in the ware systems. The maximum tuber number ( $TN_{15}$ ) observed was similar between the ware and seed systems, but there was a greater range in mean  $TN_{15}$  in ware systems with some fields having low  $TN_{15}$  values (<400). The lower CV and range in  $S_D$  in the seed systems did not translate to a lower CV range in  $TN_{15}$ , with similar  $TN_{15}$  CV ranges (Tables 2 and 3) observed in both systems. In contrast, the  $MNR_{15}$  CV ranges followed the  $S_D$  CV values in both systems with more variation in the ware fields (higher CV values and larger range of CV values).

It was hypothesised that some of the variables, particularly quality variables, may exhibit skewed distributions. However, at an individual field basis this was not observed (most variables in the range  $-1 < \gamma_1 < 1$  – data not shown) and when aggregated only exhibited moderate skewness ( $-2 < \gamma_1 < 2$  – Tables 2 and 3), which is likely to be inflated by differences in varieties and in production practices between fields.

<<INSERT TABLE 2 NEAR HERE>>

<<INSERT TABLE 3 NEAR HERE>>

### *Spatial Statistics – Manual Dig Data*

Tables 4 and 5 show the ‘average’ variogram parameters for the manually measured variables within the ware and seed systems respectively. The tables also show the derived indicators of spatial variation (CI, MCD) as well as the conventional (non-spatial) CV. In this case the CV is calculated using all data, not on a field basis as presented in Tables 2 and 3.



In the ware fields,  $S_D$  showed high variability (high CV) but less spatial structure (high CI, low MCD) to the  $S_D$  variation. In both ware and seed fields the spatial variation in  $\%Y_M$  yield was slightly higher than that of  $Y_T$  (higher MCD and lower CI). The  $MNR_{15}$  exhibited the strongest spatial structure in the ware fields but showed a low spatial structure in seed systems when compared to  $Y_T$  and  $Y_M$ . Similarly to the  $MNR_{15}$  results, the spatial structure in  $\%Y_M$  was inverted between the ware and seed fields, with strong spatial structure in ware and poor spatial structure in seed fields.

The spatial structure of  $TN_{15}$  in the seed fields was similar to  $Y_T$  and much greater than  $MNR_{15}$ . In contrast, in the ware fields, the spatial variance of  $MNR_{15}$  was much greater (lower CI, higher MCD and range) than that of its constituents ( $TN_{15}$  and  $Y_T$ ).

All the yield production variables exhibited variogram ranges ( $a$ ) that would indicate that they could be managed with typical farm machinery, assuming swathings operations of 24-36 m, which is less than half the shortest range ( $a > 74.3$  m).  $S_D$  in ware fields had the shortest range, indicating less opportunity to spatially manage according to  $S_D$  with current machinery, however it does not indicate an inability to spatially manage, especially if section-control is available on machinery.

<<INSERT TABLE 4 NEAR HERE>>

<<INSERT TABLE 5 NEAR HERE>>

The standardised 'average' variograms from the measured manual dig data are shown in Figure 2 for ware (top) and seed (bottom) production systems. These visualise the information presented in Tables 4 and 5 respectively. The inverse pattern of the quantity ( $Y_T$  and  $Y_M$ ) and  $MNR_{15}$  variables between the two systems is obvious, with quality indicators ( $MNR_{15}$  and  $\%Y_M$ ) having relatively more spatial structure in the ware systems and relatively less spatial structure in the seed fields than  $Y_T$  and  $Y_M$ .

<<INSERT FIGURE 2 NEAR HERE>>

#### *Spatial Statistics – Yield sensor data*

The yield monitor data returned similar field average yield values to the manual dig data (data not shown) but exhibited a higher level of (semi-)variance than the manual dig data (Table 6). This is likely a compound effect of much higher density data with sensor noise, edge effects and short-term (<3 m) stochastic errors in the yield monitor data that are not present in the dig data. However the CI values are similar across all the fields (62-72 indicating a moderate spatial structure) (Table 6). The average CI from the yield data was higher, indicating less spatial structure, than the average  $Y_T$  CI from the manual dig data.

<<INSERT TABLE 6 NEAR HERE>>

<<INSERT FIGURE 3 NEAR HERE>>

The sensor-based and manual dig-based standardised average variograms are shown in Fig. 3 for the 5 ware fields that had coherent yield sensor data. The two variograms show a similar CI and range. The spatially denser sensor data clearly shows a nested structure to the variance (2 separate responses) with an initial short range effect ( $a_1 = 12.1$  m) followed by a longer effect ( $a_2 = 140.2$  m) (Table 6 and Fig. 3). This indicates that there is a large amount of variation over short ranges in the sensor yield data and the percentage of variation described by the longer range variation ( $C_{1b}/(C_0+C_{1a}+C_{1b})$ ) is small (20.3%). The total range of the yield 'average' variograms from the manual dig and yield sensor was very similar (133.5 vs. 140.2 m), indicating that both approaches have identified a similar spatial structure to the autocorrelation in the total yield data. The manual dig data did not indicate any nested effects for individual fields (data not shown) or for average  $Y_T$  variograms (Tables 4-5 and Figures 1-2), however the data density (100 sites per field) in the manual UK survey is unlikely to be of sufficient spatial density to identify nested structures. Given this limitation the two standardised yield variograms from this study demonstrated a similar trend (similar CI and range) even though they are generated from different data.

The CI values from the yield sensor are similar to those observed in other reported studies (Cambouris et al. 2006 and Whelan and Mulcahy 2016). However the variogram range in this UK survey data is longer than that reported in Australia and Canada, indicating more structured spatial variability in the UK ware production systems. It may be that the earlier studies have only accounted for the first component of nested variation as their reported ranges are quite short (< 40 m). However, it must be stressed that this yield sensor-based information is preliminary data and from a very small survey.

## Discussion

This is the first publication to provide explicit information on the spatial variability in potato quality and quantity in both ware and seed production systems and how this will impact potential site-specific management of potatoes. It provides a benchmark to help inform future spatial applications in potato, especially for those with a UK focus. It was not the intent here to explore in detail further spatial interactions between the production variables and management and environmental effects. In particular the data set is unique in providing the first spatial information related to seed production systems and providing information on tuber size distribution and marketable yield, instead of just total yield.

The high variance (CV) but low spatial structure in  $S_D$  in the ware systems is perhaps an area for growers to address. The cause of this is unclear. Stem density was used as a surrogate for canopy size in the study and it is recognised that this has some limitations. Stem density is known to be determined by seed size and the physiological status of the seed, which in turn is dependent on planting date and conditions, the environmental growing condition of the seed crop and the subsequent storage of the seed potatoes. The care taken in using the same seed potato batch across fields in a given year and never mixing batches within any field was done to minimise this potential effect. Ideally stem density should also include information related to the level of branching and canopy size and shape to be properly interpreted, but such information was not possible. Recorded plant numbers were very

variable in ware fields with between 4-11 plants per 3 m dig recorded. The plant density variation comes from two potential sources; firstly the determination of plant numbers at harvest was not always easy, which was why stems were preferred as a more accurate count and reflection of canopy, and secondly, it was hypothesised that the belt-planters used in the ware systems were introducing a large variability in planting density, which, with relatively uniform seed quality and age, will influence local stem densities. The variable planting density, and subsequent plant density, masked at least some of the crop variation that may be driven by environmental and edaphic variation. If more uniform plant and stem densities are desirable, it would seem that improved planters are the first step required to achieve this. It is an edict in precision agriculture that it is important to be performing correct or optimum 'uniform' management before moving to differential management (Whelan and Taylor 2013 p8). As plant density is known to affect both yield and tuber size distribution (Zheng et al. 2016) it would seem desirable to have better control of this factor in the system.

In contrast, in the seed production fields, albeit constrained to one operator in one region,  $S_D$  showed less total variance (lower CV) than the other yield components ( $Y_T$ ,  $Y_M$ ), but the spatial structure was similar. In these fields, spatial variance in  $S_D$  appears to be driving yield and spatial  $S_D$  information should be more relevant as an input in the context of spatial yields models or field operations. The cause of the lower  $S_D$  variance in the seed fields, compared to the ware fields (Tables 2 and 3), is hypothesised to be due to the use of a cup-planter in the seed system (not a belt planter as used in the ware fields) and/or the closer planting density in seed systems. What is clear from the data is that the less variable stem density reduces the stochastic (and management) variance in the data and allows for environmental and/or edaphic variability to be more strongly expressed in crop variability.

The variogram profiles presented in Figure 2 and statistics in Table 4 and 5 clearly indicated two contrasting opportunities between the seed and ware production systems. There seems a greater opportunity to manage quality, particularly the  $MNR_{15}$  variable (average tuber weight) in the ware fields compared to yield. This could be addressed in two ways – either altering the length of the growing season, in this case prolonging it to ensure sufficient dry matter accumulation to grow the observed (actual)  $TN_{15}$  to marketable size, or by trying to alter  $TN_{15}$  (via variable planting and/or irrigation) so that accumulated dry matter production at a site can be used to grow a higher percentage of tubers into the marketable size range. In these ware fields a uniform burn-down operation was used. Given this, and the observed variation in  $Y_T$  and  $TN_{15}$ , it is not surprising that a high spatial variance in  $\%Y_M$  was also observed. An understanding mid-season of the spatial variance in  $MNR_{15}$  within a field, through targeted test digs and using field tools, such as the PotatoSize™ app (James Hutton Institute 2017), may provide an indication of likely  $\%Y_M$ . This would provide a potential opportunity for mid/late-season differential canopy management (including senescence) to target the highest possible local  $\%Y_M$ . For example, halting areas with advanced tuber development, and a risk of oversized tubers, using variable-rate (date) haulm destruction strategies may be an option to allow slower developing areas to 'catch-up' and to achieve a more uniform tuber size distribution and higher overall field  $\%Y_M$ .

In the seed systems it appears that managing  $Y_T$  presents the greatest opportunity for spatial management and improved profitability. In contrast, the  $\%Y_M$  and  $MNR_{15}$  spatial variance is less and more unstructured. However, given that the  $\%Y_M$  was between 86-96% (Table 3) in the fields, this result is not surprising. The worst  $\%Y_M$  in the seed fields was close to the best  $\%Y_M$  in the ware fields (86% vs. 89% Tables 2 and 3). It was observed that there was a low spatial variance in  $MNR_{15}$ , which if

the crop is stopped at the correct time and the shorter-season crop has not had enough time to affect spatial differential tuber growth rates, should (and did) equate to a low variance in  $\%Y_M$  in seed crops. The opportunity for improved productivity in the seed fields will therefore come from increasing tuber number (and therefore potential yield) in areas with lower tuber initiation and ensuring that the local seasonal dry matter accumulation also increases to match the higher sink demand for carbohydrate. The drivers for the strong spatial variance in tuber number in the seed fields are not clear from this survey and require further investigation of the environment and the agronomy of production within the fields. It is likely that different drivers will occur in different fields and require a field-specific agronomic solution. It appears that the spatial variance in  $TN_{15}$  and  $Y_T$  interacts to dampen the spatial variance in  $MNR_{15}$  in the seed fields, while they interact to amplify the spatial variance in  $MNR_{15}$  in the ware fields.

While it appears opportunities exist to manipulate quality or quantity in potato fields, attempting to independently manipulate  $TN_{15}$ ,  $MNR_{15}$  or  $Y_T$ , without considering interactions is likely to create problems. Correlations at a field-scale between  $TN_{15}$  and  $Y_T$  ranged from 0.24 - 0.79 for seed systems and 0.22 - 0.83 for ware systems. Adjusting yield goals therefore influences  $TN_{15}$  (and vice versa) and will have some effect on  $MNR_{15}$ . Similarly, a variable rate planting is employed in ware fields at the start of the season to try to spatially manage  $TN_{15}$ , care must be taken in managing canopy size (dry matter accumulation) so that it is not also linearly affected. If this occurred, the gains in managing  $TN_{15}$  could be lost. The different variogram response between  $S_D$  (an indicator of canopy size) and  $TN_{15}$  and  $MNR_{15}$  (Fig. 2a) in ware fields indicates that this was not likely in these data. A detailed analysis of such interactions are beyond the aim of this paper, but an area that needs to be considered and further researched if differential management is adopted. Such interactions are likely to be field or site-specific.

The high spatial density yield monitor data indicated that there were nested spatial structures in the yield data in 4 of the 5 fields (although only 3 had a best fit with a double model) and in the average variogram (Table 6). The first question is – are these nested structures real or a sensor artefact? The average variogram, derived from five fields, from two locations using two different harvesters and sensors, exhibited this nested behaviour, so it seems unlikely to be just a local yield monitor effect. The first range ( $a_1$ ) is in the order of 12-20 m. Management operations are variable but tend to be done at ~24 m swaths, indicating that this short-range variation may be driven at least in part by in-season management. Nested variance structures in the quality data, such as  $MNR_{15}$ , would support a management effect, however a more intensive local survey would be needed to determine this. The longer range variation observed in the manual dig data and the second nested component in the yield data indicate other environmental factors aside from management are at play.

From the ‘average’ yield monitor variogram derived from 5 UK fields, 44% of the yield variance has some spatial structure ( $CI = 55.93$ ), with approximately half as short-range variance ( $c_{1a}$ ) and half as long-range variance ( $c_{1b}$ ) (Table 6). If the short-range variance is predominantly management driven, then approximately half the observed spatial (manageable) yield variance is created by management operations. Only 22% of the observed yield variation is driven by longer-range factors that are likely to be edaphic or other environmental factors. Yield variance could be reduced by differential management to reflect different growing environments, but there appears to be equally as much opportunity to reduce yield variation by improving general crop operations and reducing management effects on yield without adopting differential management. It must be stressed that while these yield

monitor data conform with other reported yield monitor data, the interpretation is limited by a limited amount of data (both in the data collected in this survey and in the reported literature). The industry would benefit from having access to more yield monitor data or at least the statistics associated with more yield monitor data.

Reducing yield variance (achieving a more uniform yield) is likely to be an unreasonable goal for a producer. The objective should be to obtain a uniformly high quality of product, potentially by targeting different yield levels in different growing environments. The results of the manual survey clearly indicated that both yield and quality are spatially variable within UK fields, although there are differences between ware and seed production systems. Regardless, in both systems there is sufficient spatial structure in either quantity and/or quality attributes to consider spatial management. Having quantified the expected variance, and where it occurs, next steps in any precision potato management will of course depend on the underlying agronomy, the local drivers of variation, management and the economics of adoption. The yield variance from very high-resolution yield monitor data had a nested structure that is posited to have both a management and an environmental component to it. Given this, it is hypothesised that the 'quality' data will also exhibit a nested structure, although the density of the manually dug data collected in the survey was too sparse to reveal this.

## Conclusion

Non-spatial and spatial statistics of total yield, marketable yield, tuber size and stem density, from a systematic manual survey of 13 fields and yield monitoring in 5 fields, have been presented as a basis for future spatial analysis and management of ware and seed potato systems. The systematic survey showed that there was considerable spatial and non-spatial variability in quantity, quality and canopy (stems) in both ware and seed potato production systems in the UK. A variogram analysis indicated that the biggest opportunity for site-specific potato management in ware production systems lies in better management of tuber size distribution and improving the percentage of total yield that is marketable. For seed production systems, the biggest opportunity is in managing the total yield, rather than the marketable yield. The large amount of variance with low spatial structure observed in the manual stem density (and plant density) data within 3 m digs, together with the short range nested total yield variance observed in the yield monitor data, indicated that approximately half the observed ware production variation may be associated with management effects, particularly variability in planting.

## References

- Al-Gaadi, K.A., Hassaballa, A.A., Tola, E., Kayad, A.G., Madugundu, R., Alblewi, B., Assiri, F. 2016. Prediction of potato crop yield using precision agriculture techniques. PLoS ONE 11, art. no. e0162219. <https://doi.org/10.1371/journal.pone.0162219>
- Allaire, S.E., Cambouris, A.N., Lafond, J.A., Lange, S.F., Pelletier, B., Dutilleul, P. 2014. Spatial variability of potato tuber yield and plant nitrogen uptake related to soil properties. Agron. J. 106, 851-859

521 Cambardella, C.A., Moorman, T.B., Novak, J.M., Parkin, T.B., Karlen, D.L., Turco, R.F., Konopka A.E.  
522 1994. Field-scale variability of soil properties in central Iowa soils. *Soil Sci. Soc. Amer. J.* 58, 1501-  
523 1511.

524 Cambouris, A.N., Zebarth, B.J., Ziadi, N., Perron, I. 2014 Precision Agriculture in Potato Production.  
525 *Pot. Res.* 57, 249-262.

526 Cambouris, A.N., Nolin, M.C., Zebarth, B.J., Laverdiere, M. 2006. Soil management zones delineated  
527 by electrical conductivity to characterize spatial and temporal variations in potato yield and in soil  
528 properties. *Am. J. Pot. Res.* 83, 381-395

529 Chen, H., Leinonen, I., Marshall, B., Taylor, J.A. 2017. Conceptual Spatial Crop Modelling of Potato  
530 Production. *Adv. Anim. Biosci. (ECPA)* 2017, 8, 678-683.

531 Davenport, J., Hattendorf, M. 2000. Using Site-specific Approaches to Advance Potato Management  
532 in Irrigated Systems. *HortTech.* 10, 452-457.

533 Davenport, J.R., Redulla, C.A., Hattendorf, M.J., Evans, R.G. and Boydston, R.A. (2002). Potato Yield  
534 Monitoring on Commercial Fields. *HortTechnology.* 12, 289-296.

535 Finke, P.A. 1992. Integration of remote sensing data in the simulation of spatially variable yield of  
536 potatoes. *Soil Technol.* 5, 257-270.

537 Han, S., Hummel, J.W., Goering, C.E., Cahn, M.D. 1994. Cell size selection for site-specific crop  
538 management. *Trans. Am. Soc. Agric. Eng.* 37, 19-26.

539 Hess, J.R., Svoboda, J.M., Hoskinson, R.L., Hempstead, D.W., Jones, W.B. 1998. Spatial potato seed  
540 piece placement monitoring system. In P.C. Robert (ed.). *Proc 4th Int. Conf. Precis. Agric.* ASA–CSSA–  
541 SSSA Press, Madison, WI, 653-660

542 Hetzroni, A., Zig, U., Warshavsky, S., Yosef, S. 2007. Quality mapping of field crops. In J.V. Stafford  
543 (ed.), *Precision Agriculture '07. Proc. 6<sup>th</sup> Eur. Conf. Precis. Agric.*, Skiathos, Greece, June 3-6, 2007.  
544 Wageningen Academic Publishers. 297-302.

545 Holmes, A., Jiang, G. 2017. Precision agriculture for New Zealand potatoes – effect of variable yield,  
546 tuber size and income. In *Proc 7<sup>th</sup> Asian-Australasian Conf. Precis. Agric.* October 16-18, 2017,  
547 Hamilton, NZ

548 James Hutton Institute (2017) The PotatoSize app. Accessed via iTunes and Google Play, 22  
549 November 2017.

550 Kempenaar, C., Been, T., Booij, J., van Evert, F., Michielsen, J-F. and Kocks, C. (2018). Advances in  
551 Variable Rate Technology Application in Potato in The Netherlands. *Potato Res.* p1-11. (Available via  
552 On-line First).

553 McBratney, A.B., Pringle, M.J. 1999. Estimating average and proportional variograms of soil  
554 properties and their potential use in precision agriculture. *Precis. Agric.* 1, 125-152.

555 Mackerron, D.K.L., Marshall, B., Jefferies, R.A. 1988. The distributions of tuber sizes in droughted and  
556 irrigated crops of potato. II. Relation between size and weight of tubers and the variability of tuber-  
557 size distributions. *Pot. Res.* 31, 279-288.

558 Minasny, B., McBratney, A.B., Whelan, B.M. 2005. VESPER version 1.62. Precision Agriculture  
559 Laboratory, The University of Sydney, NSW.  
560 <https://sydney.edu.au/agriculture/pal/software/vesper.shtml>. Accessed 22 November 2017

561 Persson, D.A. 1998. Potato yield mapping with an optical sensor. In BJ Petoskey (ed.), Proc. 1st Int.  
562 Conf. Geospatial Inf. Agric. Forest. ERIM Int., Ann Arbor, USA, 618-623.

563 Persson, A., Pilesjö, P., Eklundh, L. 2005. Spatial Influence of Topographical Factors on Yield of Potato  
564 (*Solanum tuberosum* L.) in Central Sweden. *Precis. Agric.* 6, 341. [https://doi.org/10.1007/s11119-](https://doi.org/10.1007/s11119-005-2323-6)  
565 005-2323-6

566 Pringle, M.J., McBratney, A.B., Whelan, B.M., Taylor, J.A. 2003. A preliminary approach to assessing  
567 the opportunity for site-specific crop management in a field, using a yield monitor. *Agric. Syst.* 76, 273-  
568 292.

569 Rees, H.W., Chow, T.L., Gregorich, E.G. 2007. Spatial and temporal trends in soil properties and crop  
570 yield at a site under intensive up- and down-slope potato production in northwestern New  
571 Brunswick. *Can. J. Soil Sci.* 87, 383-398. <https://doi.org/10.4141/CJSS07017>.

572 Sands P.J., Regel, P.A. 1983. V. A simple model for predicting graded yields. *Field Crops Res.* 6(C), 25-  
573 40.

574 Schneider, S.M., Boydston, R.A., Han, S., Evans, R.G. 1997. Mapping of potato yield and quality. In JV  
575 Stafford (ed.) *Precision Agriculture '97. vol. I: Spatial variability in soil and crop.* Bios Scientific Publ.,  
576 London pp 253–261.

577 Schneider, S.M., Han, S., Campbell, R.H., Evans, R.G., Rawlins, S.L. 1996. Precision agriculture for  
578 potatoes in the Pacific Northwest. In PC Robert (ed.) *Proc. 3rd Intl. Conf. Precis. Agric.* ASA–CSSA–  
579 SSSA Press, Madison, Wisconsin. pp 443–452.

580 Starr, G.C. 2005. Assessing temporal stability and spatial variability of soil water patterns with  
581 implications for precision water management. *Agricultural Water Manag.* 72, 223-243.

582 Taylor, J.A., McBratney, A.B., Whelan, B.M. 2007. Establishing management classes for broadacre  
583 grain production. *Agron. J.* 99, 1366-1376. <https://doi.org/10.2134/agronj2007.0070>.

584 Tisseyre, B., McBratney, A.B. 2008. A technical opportunity index based on mathematical  
585 morphology for site-specific management: an application to viticulture. *Precis. Agric.* 9, 101-113.  
586 <https://doi.org/10.1007/s11119-008-9053-5>.

587 van Evert, F.K., Gaitán-Cremaschi, D., Fountas, S., and Kempenaar, C. (2017). Can Precision  
588 Agriculture Increase the Profitability and Sustainability of the Production of Potatoes and Olives?  
589 *Sustainability* 9, ISSN 2071-1050.

590 Van Uffelen, C.G.R., Verhagen, J., Bouma, J. 1997. Comparison of simulated crop yield patterns for  
591 site-specific management. *Agric. Syst.* 54, 207–222.

592 Verhagen, A., Booltink, H.W.G., Bouma, J. 1997. Site-specific Management: Balancing Production  
593 and Environmental Requirements at Farm Level. *Agric. Syst.* 49, 369-384.

594 Verhagen, J. 1997. Site specific fertiliser application for potato production and effects on N-leaching  
595 using dynamic simulation modelling. *Agric. Ecosyst. Environ.* 66, 165-175.

596 Webster, R., Oliver, M.A. 1992. Sample adequately to estimate variograms of soil properties. J. Soil  
597 Sci. 43, 177–192.

598 Whelan, B., Mulcahy, F. 2016. Understanding spatial variability in potato cropping to improve yield  
599 and production efficiency. Project Report PT13000. Horticulture Australia Limited.  
600 <https://ausveg.com.au/app/data/technical-insights/docs/PT13000.pdf> Last accessed 30/11/2017.

601 Whelan, B., Mulcahy, F. 2017. A strategy to instigate SSCM in Australian potato production. Adv.  
602 Anim. Biosci. 8, 743-748. <https://doi.org/10.1017/S2040470017000401>

603 Whelan, B.W., Taylor, J.A. 2013. Precision Agriculture for Grain Production Systems. CSIRO Publishing,  
604 Dickson, ACT, Australia (8 Chapters; 208p) ISBN:9780643107472.

605 Zheng, S-L., Wang, L-J., Wan, N-X., Zhong, L., Zhou, S-M., He, W., Yuan, J-C. 2016. Response of Potato  
606 Tuber Number and Spatial Distribution to Plant Density in Different Growing Seasons in Southwest  
607 China. Front Plant Sci. 7, Article 365. <https://doi.org/10.3389/fpls.2016.00365> ISSN:1664-462X.



608 **Table 1** Description of the fields used in the survey

Field	Area (ha)	County	Variety	Marketable size grades (mm)	Riddle Grading (mm)	Year
Field 1	2.7	Staffordshire	Russet Burbank	>50	5	2015
Field 2	13.6	Staffordshire	Russet Burbank	>50	5	2015
Field 3	11.6	Yorkshire	Saxon	>50	10	2015
Field 4	7.6	Yorkshire	Maris Piper	>50	10	2015
Field 5	3.8	Staffordshire	Innovator	>50	5	2015
Field 6	7.4	Angus	Picasso (seed)	35 – 65	10	2015
Field 7	5.7	Angus	Markies (seed)	35 – 65	10	2015
Field 8	6.4	Angus	Markies (seed)	35 – 65	10	2015
Field 9	9.1	Staffordshire	Russet Burbank	>50	5	2016
Field 10	7.3	Yorkshire	Pentland Dell	>50	5	2016
Field 11	6.1	Yorkshire	Russet Burbank	>50	5	2016
Field 12	11.2	Angus	Markies (seed)	35 – 65	5	2016
Field 13	5.8	Angus	Picasso (seed)	35 – 65	5	2016

609

610 **Table 2** Non-spatial statistics ranges from the ware potato production fields.

Variable	Mean ( $\mu$ )	Standard Deviation ( $\sigma$ )	Coefficient of Variation (CV) (%)	Skewness ( $\gamma_1$ )	Range
Stems ( $\times 10^3 \text{ ha}^{-1}$ )	67.6 - 129.9	8.3 - 37.5	12.4 - 31.5	-0.05 - 1.33	37.0 – 240.7
Total Yield ( $Y_T$ ) (Mg $\text{ha}^{-1}$ )	44.8 - 93.8	6.5 - 14.1	11.0 - 17.9	-0.49 - 0.41	32.9 - 65.3
Marketable Yield ( $Y_M$ ) (Mg $\text{ha}^{-1}$ )	28.8 - 81.8	6.3 - 16.6	14.2 - 29.7	-0.37 - 0.43	32.5 – 76.0
Percent Marketable Yield ( $\%Y_M$ ) (%)	63.2 - 88.6	3.9 - 12.8	4.5 - 20.1	-0.91 - 0.51	19.3 - 69.6
Tuber No. (TN) ( $\times 10^3$ $\text{ha}^{-1}$ )	181.3 - 581.6	37.5 – 171.0	11.1 - 29.4	-0.46 - 0.93	170.4 - 698.2
Mass:Number Ratio ( $\text{MNR}_{15}$ ) (g tuber $^{-1}$ )	84.0 - 340.5	16.7 – 67.9	11.4 - 33.5	-0.07 - 1.40	44.3 – 501.1

611

612

613 **Table 3** Non-spatial statistics ranges from the seed potato production fields.

Variable	Mean ( $\mu$ )	Standard Deviation ( $\sigma$ )	Coefficient of Variation (CV)	Skewness ( $\gamma_1$ )	Range
Stems ( $\times 10^3 \text{ ha}^{-1}$ )	260.7 – 369.9	42.9 - 67.4	14.1 - 18.2	-0.50 - 0.25	148.2 - 544.4
Total Yield ( $Y_T$ ) (Mg $\text{ha}^{-1}$ )	23.7 - 56.6	5.4 - 8.9	11.0 - 22.6	-0.50 - 0.34	26.1 - 40.3
Marketable Yield ( $Y_M$ ) (Mg $\text{ha}^{-1}$ )	20.9 - 54.1	5.4 - 8.5	10.5 - 25.6	-0.5 - 0.29	25.1 - 38.3
Percent Marketable Yield ( $\%Y_M$ ) (%)	86.5 - 96.2	3.0 - 8.4	3.1 - 9.7	-1.59 - -0.36	11.8 - 35.5
Tuber No. (TN) ( $\times 10^3$ $\text{ha}^{-1}$ )	547.6 – 637.1	73.7 - 147.4	13.5 - 26.9	-0.52 - 0.41	418.5 - 703.7
Mass:Number Ratio ( $\text{MNR}_{15}$ ) (g tuber $^{-1}$ )	37.2 – 101.5	6.4 - 18.7	12.9 - 18.8	0.26 - 0.88	23.3 - 149.3

614

615

**Table 4** 'Average' variogram parameters and derived spatial statistics for the ware production system fields

Variable	$c_0^*$	$c_1^*$	$a$ (m)	Model	CI	MCD	CV
Stems ( $\times 10^3 \text{ ha}^{-1}$ )	320.5	131.8	48.3	Spherical	70.9	6.4	31.6
Total Yield ( $Y_T$ ) ( $\text{Mg ha}^{-1}$ )	49.2	43.8	123.6	Spherical	53.3	21.7	28.7
Marketable Yield ( $Y_M$ ) ( $\text{Mg ha}^{-1}$ )	45.7	58.6	113.8	Spherical	43.8	24.0	39.3
Percent Marketable Yield ( $\%Y_M$ ) (%)	44.1	50.8	206.9	Spherical	46.5	41.5	16.5
Tuber No. (TN) ( $\times 10^3 \text{ ha}^{-1}$ )	3795.0	2743.7	195.5	Spherical	58.0	30.8	41.1
Mass:Number Ratio ( $\text{MNR}_{15}$ ) ( $\text{g tuber}^{-1}$ )	552.2	901.2	262.6	Spherical	38.0	61.1	45.6

\* the nugget ( $c_0$ ) and sill ( $c_1$ ) variances expressed in the units measured.

**Table 5** ‘Average’ variogram parameters and derived spatial statistics for the seed production system fields

Variable	$c_0^*$	$c_1^*$	$a$ (m)	Model	CI	MCD	CV
Stems ( $\times 10^3 \text{ ha}^{-1}$ )	1544.1	1276.6	257.9	Spherical	54.7	43.8	21.9
Total Yield ( $Y_T$ ) ( $\text{Mg ha}^{-1}$ )	34.0	26.0	289.0	Spherical	56.6	47.0	32.8
Marketable Yield ( $Y_M$ ) ( $\text{Mg ha}^{-1}$ )	24.8	26.3	312.0	Spherical	48.6	60.1	32.7
Percent Marketable Yield ( $\%Y_M$ ) (%)	16.5	8.9	74.3	Spherical	64.9	9.8	7.2
Mass:Number Ratio ( $\text{MNR}_{15}$ ) ( $\text{g tuber}^{-1}$ )	117.7	85.3	84.4	Spherical	58.0	13.3	35.4
Tuber No. ( $1000\text{s ha}^{-1}$ )	7353.2	6419.1	290.2	Spherical	53.4	50.7	19.4

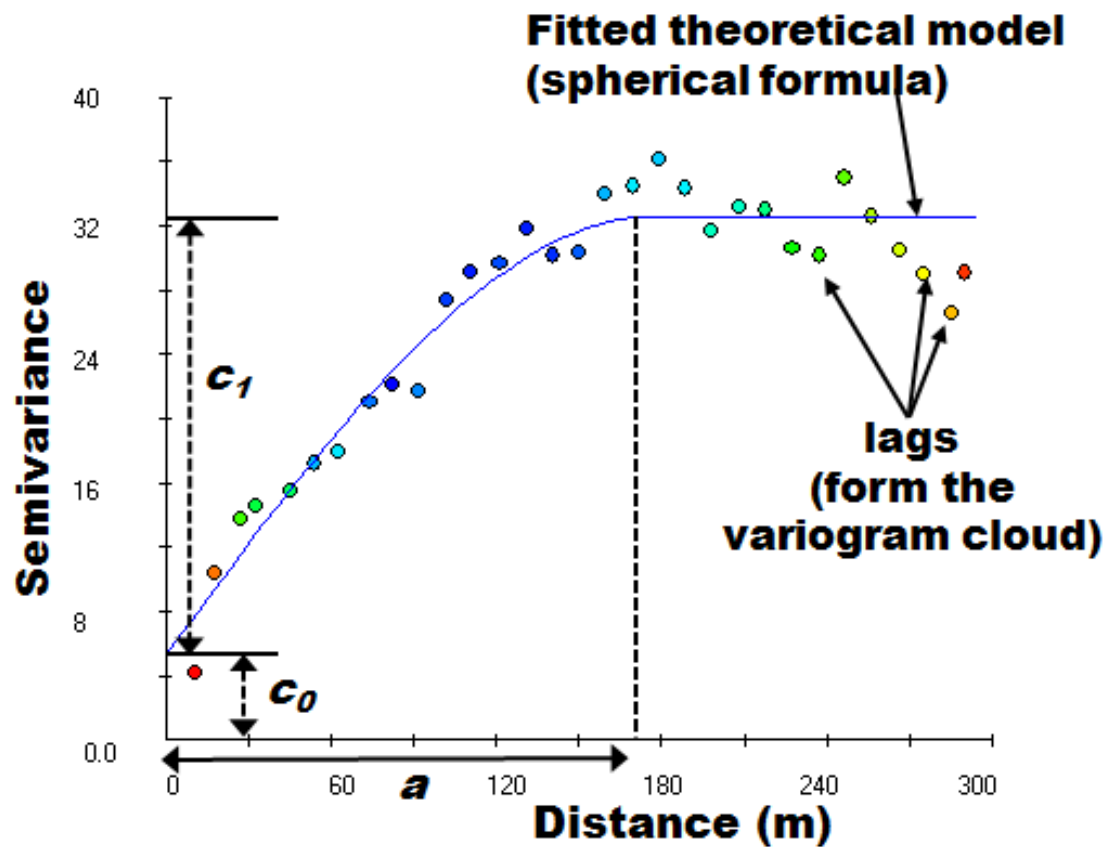
\* the nugget ( $c_0$ ) and sill ( $c_1$ ) variances expressed in the units measured.

**Table 6** Individual field and ‘average’ yield variogram parameters and statistics derived from yield data collected by an on-harvester yield monitor and the equivalent ‘average’ yield variogram and statistics from manual digs (3 m plots) with all yield data expressed as Mg ha<sup>-1</sup>.

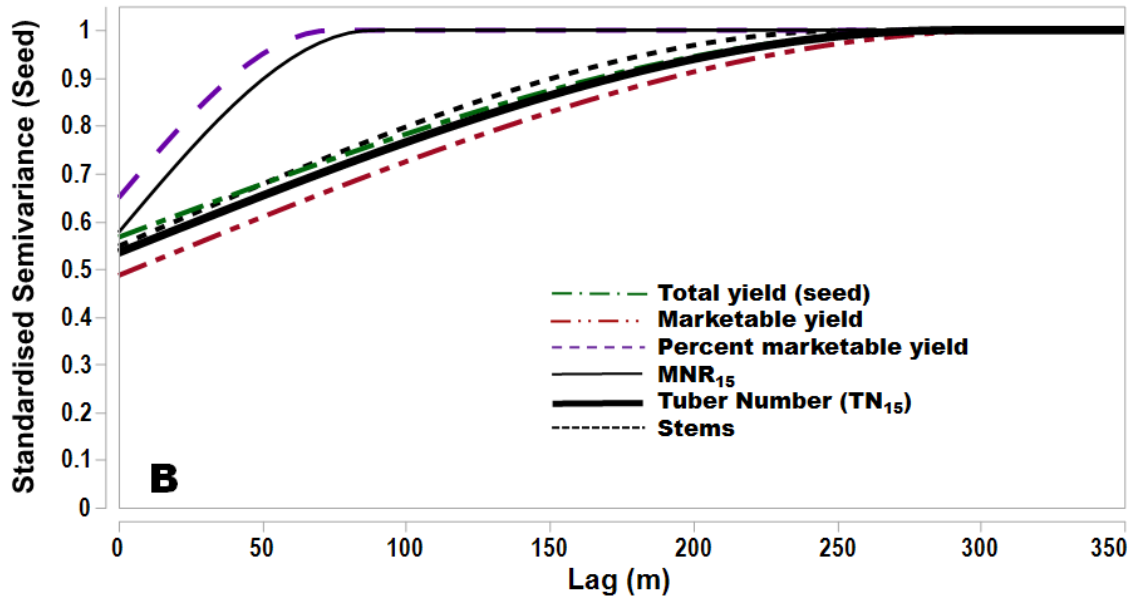
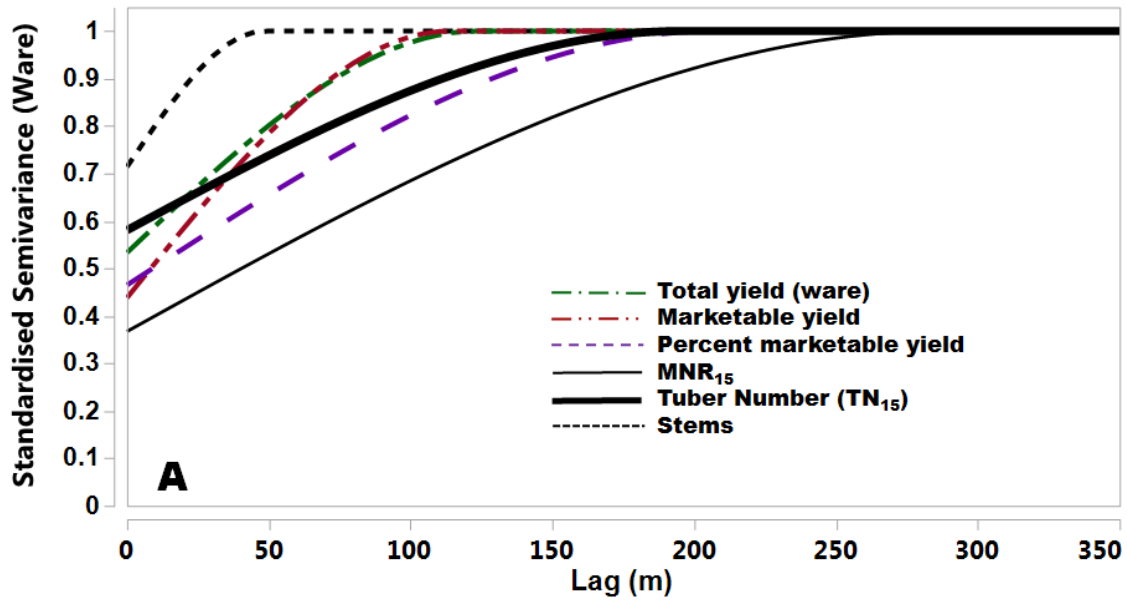
Yield Monitor Data	$c_0^*$	$c_{1a}^*$	$a_1$ (m)	$c_{1b}^*$	$a_2$ (m)	Model	CI	MCD
Field 1 <sup>§</sup>	123.3	50.6	425.4			Spherical	70.9	46.4
Field 2	226.2	118.4	140.5			Spherical	65.6	18.1
Field 5	249.5	64.6	19.8	153.9	224.4	Spherical	53.3	31.4
Field 10	433.2	178.4	14.3	267.6	333.2	Double Spherical	49.3	16.2
Field 11	144.6	29.1	17.3	112.8	370.0	Double Spherical	50.5	4.5
†^Average yield sensor	297.2	115.5	12.1	118.7	140.2	Double Spherical	55.9	25.2
‡Average Manual Digs	56.02	62.4	133.5			Spherical	47.3	26.4

† CI and MCD calculate using summed  $c_{1a}$  and  $c_{1b}$  and  $a_1$  and  $a_2$  values; ^ All fields randomly subset to 10,000 points before merging the subset data to calculate the ‘average’ yield variogram. ‡ Calculated only for the same fields as yield sensor data was available for in contrast to  $Y_T$  in Table 4 that was calculated with all available data.

<sup>§</sup>Although Field 1 best fit was a single spherical model there was a very short range component ( $a = 6.4m$ ) observed as well. Field 2 had no nested structure. \* the nugget ( $c_0$ ) and sill ( $c_1$  and  $c_{1b}$ ) variances expressed in the units measured.

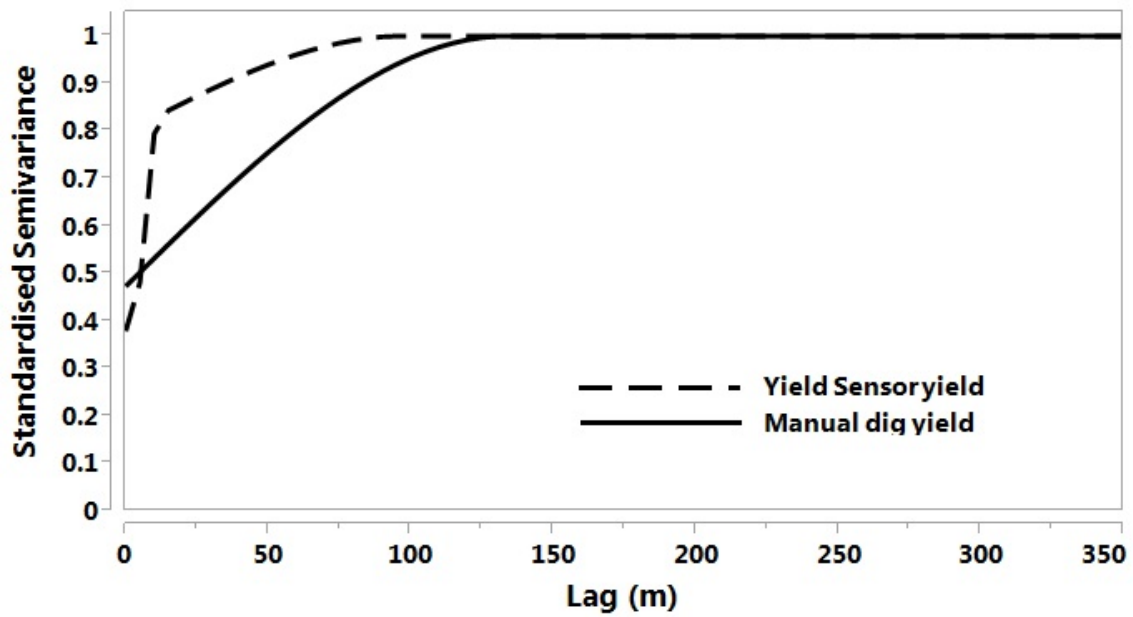


**Figure 1:** Example of a variogram showing the ‘experimental variogram’ (or variogram cloud) composed of lags that represent the average semi-variance of points separated by a certain distances, a mathematical model fitted to the points that is known as the theoretical variogram, and an illustration of the parameters (nugget variance ( $c_0$ ), sill variance ( $c_1$ ) and range ( $a$ )) that are used to describe the mathematical model.



**Figure 2:** Standardised 'average' variograms of quantity metrics (Total yield, Marketable yield and Percent marketable yield), quality metrics (ratio of mass to number of tubers ( $MNR_{15}$ )) and a crop growth metric (Stems counts) in both ware (A) and seed (B) potato production systems. Units as listed in Tables 1-4.





**Figure 3:** Standardised 'average' yield variogram models derived from yield sensor (dashed) and manual dig (solid line) data for ware production systems in the UK. Both models were derived using the same subset of fields i.e. only fields where both yield sensor and manual dig data were available.